Docket No.: 4555-105 US

I HEREBY CERTIFY THAT THIS CORRESPONDENCE IS BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE AS FIRST CLASS MAIL IN AN ENVELOPE ADDRESSED TO: ASSISTANT COMMISSIONER OF PATENTS, WASHINGTON, D.C. 20231 ON THE DATE INDICATED BELOW.

BY: Valerie Balfour DATE: 9/11/01

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re: Application Suzanne Walker, *et al.*	: Group Art Unit: TBA
Appln. No.: 09/829,275	: Examiner: TBA
Filed: April 9, 2001	:
For: CRYSTALS OF THE *ESCHERICHIA COLI* MEMBRANE-ASSOCIATED GLYCOSYLTRANSFERASE (MurG) PROTEIN, ATOMIC COORDINATES AND THREE DIMENSIONAL STRUCTURES THEREOF, ATOMIC COORDINATES AND THREE DIMENSIONAL STRUCTURES OF BINDING DOMAINS THEREOF, IMAGES THEREOF, AND METHODS FOR CRYSTALLIZING MurG PROTEINS, MODELS OF UDP-GLYCOSYLTRANSFERASES, MurG PROTEINS AND BINDING SITES, METHODS OF MAKING MODELS, METHODS OF USING MODELS OF MurG COMPOUNDS THAT BIND, INHIBIT OR STIMULATE MurG PROTEINS, AND THERAPEUTIC COMPOSITIONS THEREOF	:

Assistant Commissioner of Patents
Washington, D.C. 20231

SIR:

PRELIMINARY AMENDMENT

Kindly enter the following amendments prior to examining this application.

In the specification:

On page 2, line 28, replace "?-linked" with "β-linked".

On page 8, line 2, replace "?-glucosyltransferase" with "β-glucosyltransferase".

On page 8, line 3, replace "?-strands are magenta, the aligned ?-helices" with "β-strands are magenta, the aligned α-helices".

On page 8, line 4, replace "?-glucosyltransferase" with "β-glucosyltransferase".

On page 215, lines 6 to 17 Delete the paragraph on page 215 starting from "Abstract" on line 6 and deleting all text to and including "glucosyltransferase structure prediction" on line 17.

On page 215, line 30, please replace "?=64.294, ?=83.520, ?????????????" with "α = 64.294, β = 83.520, γ = 65.448"

On page 217, line 17, please replace "I/?" with "I/σ"

On page 217, line 28, please replace "|F| > 2?" with "|F| >2σ"

On page 217, line 39, in footnote 1, please replace "?| I_i- <I> |/?I_i" with "Σ| I_i- <I> |/ΣI_i"

On page 217, line 40, in footnote 2, please replace "Mean isomorphous difference = ? $|F_{PH} - F_P| / ?F_{PH}$ " with "Mean isomorphous difference = $\Sigma |F_{PH} - F_P| / \Sigma F_{PH}$ ".

On page 217, line 46, in footnote 5, please replace "R-factor = ? $||F_{obs}| - |F_{calc}|| / ? |F_{obs}|$" with " R-factor = $\Sigma ||F_{obs}| - |F_{calc}|| / \Sigma |F_{obs}|$"

On page 218, lines 2, please replace "?/?" with "α/β".

On page 218, line 3, please replace "C?" with "Cα".

On page 218, line 4, please replace "?-strands and six ?-helices" with "β-strands and six α-helices".

On page 218, line 5-6, please replace "six parallel ?-strands and eight ?-helices" with "six parallel β-strands and eight α-helices".

On page 218, line 6, please replace "(?-link)" with "(α-link)".

On page 218, line 7, please replace "first ?-strand" with "first β-strand".

On page 218, line 7, please replace "The ?-strands" with "The β-strands" .

On page 218, line 8, please replace "seventh ?-strand" with "seventh β-strand".

On page 218, line 9, please replace "the ?-link of the C-domain" with "the α-link of the C-domain".

On page 218, line 12-13, please replace "from C-?5 to the loop connecting N-?5 to N-?5" with "from C-α5 to the loop connecting N-β5 to N-α5".

On page 218, line 14, please replace "The ?/? open-sheet motif" with "The α/β open-sheet motif".

On page 218, line 17-18, "please replace "one ?-strand and the amino terminus of the adjacent ?-helix" with "please replace "one β-strand and the amino terminus of the adjacent α-helix" .

On page 218, line 31, please replace "between N-?1/N-?1 and N-?4/N-?4)" with "between N-β1/N-α1 and N-β4/N-α4)".

On page 218, line 32, please replace "between C-?1/C-?1" with "between C-β1/C-α1".

On page 219, line 8, please replace "C? atoms" with "Cα atoms" .

On page 219, line 9, please replace "T4 ?-glucosyltransferase" with "T4 β-glucosyltransferase".

On page 219, line 15-16, please replace "the first ?-strand" with "the first β-strand".

On page 219, line 25, please replace "helix C-?4" with "helix C-α4".

On page 220, line 15, please replace "from the end of N-?5 to the middle of N-?5" with "from the end of N-β5 to the middle of N-α5".

On page 221, line 17, please replace "?/? supersecondary" with "α/β supersecondary".

REMARKS

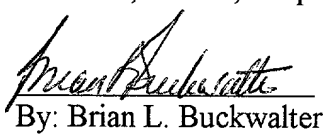
The amendments replace question marks temporarily inserted by a typist unfamiliar with Greek characters used describing amino acid or protein structure and statistical symbols and inadvertently left in the application. Their meaning is invariably obvious since "helices" refers to "α-helices" and "strands" refers to "β-strands". References to the location of specific domains, for example on page 218, line 31 is clearly depicted in Figure 3 which was submitted with the original application. References to Root Mean Square Deviation particular atoms, e.g., C-?4 on page 219 line 25, always refer to the backbone sequence which includes the α-carbon. References to the supersecondary structure, e.g., page 221, line 17, "?/?" can be found in non-symbolic form in the original text on page 4, line 32 which discusses the "alpha/beta folding pattern". The symbols for standard deviation, σ, and summation, Σ, are used in a standard format for reporting crystallography data and their meaning is obvious to the skilled crystallographer. The missing unit cell data from page line 30 can be found on page 14, line 16 of the text.

These amendments, therefore, merely correct typographical omissions in the original application; there meaning is clear from other portions of the text or drawings or from standard usage and do not represent new matter.

CONCLUSION

A clean copy of the amended pages is attached in the appendix. Applicants believe the application is presently in appropriate form for examination and a prompt Office Action on the merits of the application is respectfully requested.

Respectfully submitted,
Mathews, Collins, Shepherd & Gould, P.A.



By: Brian L. Buckwalter
Registration No. 46,585
Agent for Applicants

September 11, 2001
100 Thanet Circle, Suite 306
Princeton, NJ 08540-3662
Telephone 609-924-8555
Telecopier 609-924-3036

APPENDIX
CLEAN COPY OF AMENDED PORTIONS OF THE SPECIFICATION

Page 2, lines 26-31

MurG is the last enzyme involved in the intracellular phase of peptidoglycan synthesis (Bugg & Walsh, 1993). It catalyzes the transfer of N-acetyl glucosamine (NAG) from UDP to the C4 hydroxyl of a lipid-linked N-acetylmuramoyl pentapeptide (NAM) to form a β-linked NAG-NAM disaccharide that is transported across the cell membrane where it is polymerized and cross-linked (Fig. 1). In bacterial cells MurG associates with the cytoplasmic surface of the membrane (Bupp & van Heijenoort, 1993). However, we have found that *E. coli* MurG can be solubilized at high concentrations in active form (Ha et al., 1999).

Page 8, lines 1-11

FIG. 4. Structural analysis of the substrate binding pockets in MurG. **A**. Structural comparison between the C-terminal domain of phage T4 β-glucosyltransferase (left) and the C-terminal domain of E. coli MurG (right). The aligned six β-strands are magenta, the aligned α-helices are orange, and the other structural elements are blue. In β-glucosyltransferase, key residues involved in UDP binding are highlighted in yellow. The analogous residues in MurG are also highlighted in yellow. **B**. A close-up view of the proposed donor binding pocket in the MurG C domain with the docked UDP-GlcNAc. Conserved residues in MurG are colored magenta. The carbonyl oxygen of residue 1245 is shown in red, and its backbone nitrogen is shown in blue. **C**. The surface of *E. coli* MurG. The G loops and other conserved residues in MurG are colored magenta. The proposed membrane binding interface is also highlighted with hydrophobic residues in yellow and positively charged residues in blue.

This example describes the crystallization of the *E. coli* MurG protein and the determination of the coordinates of the three-dimensional crystal structure. This example also describes the identification of the donor nucleotide binding site, the acceptor binding site and the membrane association site of the MurG protein.

Methods

Crystallization

E. coli MurG containing a C-terminal LEHHHHHH sequence was purified as described (Ha *et al.*, 1999) and concentrated to 10 mg ml^{-1} in 20 mM Tris-HCl, pH 7.9/ 150 mM NaCl/ 50 mM EDTA. The protein concentrate was mixed with UDP-GlcNAc in a 1:3 molar ratio. Crystals were grown at room temperature using the hanging-drop vapor-diffusion method by mixing equal volumes of protein with reservoir solution (0.1 M NaMES, pH 6.5/ 0.96 M $(NH_4)_2SO_4$/0.4% Triton X-100/ 10 mM DTT). Triclinic crystals with a typical size of 0.2 mm X 0.1 mm X 0.1 mm grew within a week. The crystals belong to the P1 space group, with two molecules per asymmetric unit. The cell dimensions are a = 60.613 Å, b = 6.356 Å, = 67.902 Å, α=64.294, β=83.520, γ = 65.448.

Table 1. Summary of crystallographic and refinement data

Data set	Native	$HgCl_2$ (form A derivative)	$HgCl_2$ (form B derivative)	$(NH_4)_2WS_4$	$(NH_4)_2OsBr_6$
Resolution (Å)	1.9	2.0	1.9	2.4	2.3
Observations	288,150	101,913	245,320	44,366	106,606
Unique reflections	65,567	53,391	65,581	27,950	36,443
R_{sym}[1] (last shell)	0.032 (0.187)	0.043 (0.200)	0.042 (0.296)	0.031 (0.080)	0.056 (0.302)
I/σ (last shell)	41.9(7.0)	20.4(2.9)	29.0(3.7)	24.6(8.2)	19.6(2.5)
Completeness (78.6%) (last shell)	97.7% (96.4%)	91.4% (66.6%)	97.4% (94.0%)	83.8% (62.0%)	94.3%
MIR analysis (40.0 - 2.5 Å)					
Mean isomorphous difference[2]		0.163	0.130	0.068	0.134
Phasing power[3] (last shell)		1.09(0.73)	0.57(0.50)	0.61(0.24)	0.61(0.58)
R_{cullis} [4](last shell)		0.81 (0.91)	0.94(0.96)	0.92(0.99)	0.94(0.95)
Anomalous R_{cullis} [4](last shell)		0.96(1.00)	0.95(1.00)		
Refinement statistics					
Resolution	40.0 - 1.9 Å	R. m. s. d.[7]			
Reflections ($\|F\| > 2\sigma$)	61,989	Bonds (Å)			0.006
Protein atoms (a. u.)	5,280	Angles(°)			1.29
Water Atoms	298				
Sulfate groups	1	Ramachandran plot[8]			
R-factor-[5] 94.6%	22.0%	Residues in most favored region			
R-free[6] 5.4%	24.7%	Residues in additional allowed region			

[1] $R_{sym} = \Sigma \mid I_i - \langle I \rangle \mid / \Sigma I_i$, where I_i is the intensity of a reflection, and $\langle I \rangle$ is the average intensity of that reflection.

[2] Mean isomorphous difference = $\Sigma |F_{PH} - F_P| / \Sigma F_{PH}$, where F_{PH} and F_P are the derivative and native structure factors respectively.

[3] Phasing power is the ratio of the mean calculated derivative structure factor to the mean lack of closure error.

[4] R_{cullis} is the mean residual lack of closure error divided by the dispersive or anomalous difference.

[5] R-factor = $\Sigma |Fobs| - |Fcalc| \mid / \Sigma \mid F|$

[6] R-free is the R-factor calculated using 10% of the reflection data chosen randomly and omitted from the start of refinement.

[7] R. m. s. d., root-mean-square deviations from ideal bond lengths and bond angles.

[8] Calculated with program PROCHECK.

Page 218

The structure consists of two domains separated by a deep cleft (Fig. 2a). Both domains exhibit an α/β open-sheet structure and have high structural homology despite minimal sequence homology (RMSD = 2.02 over 85 aligned Cα atoms). The N-domain includes residues 7-163 and 341-357, and contains seven parallel β-strands and six α-helices, the last of which originates in the C-domain (Fig. 2b). The C-domain comprises residues 164-340 and contains six parallel β-strands and eight α-helices, including one irregular bipartite helix (α-link) that connects the N-domain to the first β-strand of the C-domain. The β-strands in both domains are ordered as for a typical Rossman fold. The N- and C-domains are joined by a short linker between the seventh β-strand of the N-domain and the α-link of the C-domain. This inter-domain linker and the peptide segment that joins the last helix of the C-domain to the last helix of the N-domain define the floor of the cleft between the two domains. The cleft itself is about 20 Å deep and 18 Å across at its widest point. Contacts < 4 Å across the cleft are limited primarily to interactions between residues from C-α5 to the loop connecting N-β5 to N-α5.

The α/β open-sheet motif (Rossman fold) adopted by both the N- and C-domains of MurG is characteristic of domains that bind nucleotides (Branden & Tooze, 1998). Classical Rossman domains typically contain at least one conserved glycine rich motif, with the consensus sequence GXGXXG, located at a turn between the carboxyl end of one β-strand and the amino terminus of the adjacent α-helix (Baker et al., 1992). This motif is involved in binding the negatively charged phosphates (Carugo & Argos, 1997). There are three glycine rich loops (G loops) in *E. coli* MurG (Fig. 3a) that may be variants on the phosphate binding loops found in other dinucleotide binding proteins (see below).

Sequence homology

Amino acid sequences for eighteen MurG homologs are now available. The sequence similarity between *E. coli* MurG and homologs from other bacterial strains ranges from less than 30% to more than 90% depending on the evolutionary relationship between the organisms. In all MurG homologs, however, there are several invariant residues. Fig. 3a shows a sequence alignment for a subset of MurG homologs with the invariant and highly conserved residues indicated. These residues, which include the three G loops, have been

highlighted in the *E. coli* MurG structure (Fig. 3b). Almost all of the invariant residues are located at or near the cleft between the two domains. Two of the G loops are found in the N domain (between N-β1/N-αl and N-β4/N-α4) and one is found in the C-domain (between C-β1/C-αl). The strict conservation of the highlighted

residues among different bacterial strains, and their location as determined from the *E. coli* MurG structure, implicates them in substrate binding and catalytic activity.

Structural homology reveals the donor binding site

The three-dimensional backbone structure of *E. coli* MurG was compared to known protein structures, including the three other NDP-glycosyltransferase structures that have been reported (Chamok & Davies, 1999; Gastinel *et al.*, 1999; Vrielink *et al.*, 1994). The C-terminal domain was found to have significant structural homology (RMSD= 2.218 Å for 89 aligned Cα atoms) to the C-terminal domain of phage T4 β-glucosyltransferase (BGT), an enzyme that catalyzes the glucosylation of hydroxymethyl-cytosines in duplex DNA. A co-crystal structure of BGT with UDP bound in the C-terminal domain reveals the topology of the UDP binding pocket and also shows important contacts to the nucleotide (Moréra *et al.*, 1999; Vrielink *et al.*, 1994). These contacts include: a) hydrogen bonds from the backbone amide of 1238 to the N3 and O4 positions of the base; b) hydrogen bonds between the carboxyl side chain of E272 and the O2' and O3' hydroxyls of the ribose ring; and c) contacts from a GGS motif in the loop following the first β-strand of the C domain to the alpha phosphate of UDP. The structurally homologous C-domain of MurG contains a topologically similar pocket (Fig. 4a). Furthermore, even though the two domains share only 11% sequence identity overall, there are identical residues in the same spatial location in *E. coli* MurG and in BGT. Based on this comparison, we have concluded that the C-domain of *E. coli* MurG is the UDP-GlcNAc binding site.

We have docked UDP-GlcNAc into the C-domain of E. coli MurG using the information on how UDP binds to BGT as a guide. As shown in Figure 4b, the uracil is held in place by contacts from the N3 and O4 atoms to the backbone amide of 1245. The O2' and O3' hydroxyls on the ribose sugar are within hydrogen bonding distance of the invariant glutamate residue (E269) in the middle of helix C-α4. The conserved GGS motif in G loop 3 is positioned to contact the alpha phosphate. When these contacts are made, the UDP-GlcNAc substrate fits nicely into a pocket in the C-domain, where it is surrounded by many of the invariant residues identified through sequence analysis of other MurG homologs. It is possible to propose roles for some of these invariant residues from the model. For example, the side chain of R261 can be rotated to contact the second phosphate;

this contact may help explain why UDP binds significantly better to MurG than UMP. We propose that R261 plays an important role in catalysis by stabilizing the UDP leaving group via electrostatic interactions. The side chain of Q289 is within hydrogen bonding

Page 220, lines 4-17

The acceptor binding site

Structural considerations suggest that the primary acceptor binding site is located in the N-terminal domain of MurG. This domain contains three highly conserved regions, two of which are glycine-rich loops that face the cleft (Fig 3a and 4c). These G loops are reminiscent of the phosphate binding loops found in other nucleotide binding proteins, and are most likely involved in binding to the diphosphate on Lipid I. The N-termini of the helices following each G loop form opposite walls of a small pocket between the G loops. The helix dipoles create a positively charged electrostatic field in the pocket that can stabilize the negative charged diphosphates. When the diphosphate of the acceptor is anchored in the pocket created by the G-loops, the MurNAc sugar emerges into the cleft between domains and the C4 hydroxyl can be directed towards the anomeric carbon of the GlcNAc for attack on the face opposite the UDP leaving group. The third conserved region in the N domain spans the loop from the end of N-β5 to the middle of N-α5. Kinetic analysis of mutants is required to evaluate the roles of these residues (Ha *et al.*, 1999; Men *et al.*, 1998).

Page 22,1lines 11-25

In addition to this structural homology, we have identified a strikingly similar sequence motif in the MurG family and certain other UDP-glycosyltransferase families. This sequence motif spans about a thirty amino acid stretch in the C-domain of MurG and includes most of the invariant residues found in that domain. As shown in Figure 3a, a similar motif is found in the UDP-glucuronosyltransferases (Mackenzie, 1990). Certain residues are identical, including a number of prolines and glycines, and the spacing between them is invariant. This suggests that the UDP-glucuronosyltransferases contain a region of α/β supersecondary structure that is involved in a similar function as the corresponding region in MurG (Fig. 3c). This region binds the donor sugar. By analyzing the similarities and differences between the conserved residues in this subdomain in the MurG family and other UDP-glycosyltransferase

families, it may be possible to identify - and perhaps alter - residues that are involved in determining donor selectivity. We note that it would be useful to be able to manipulate donor specificity because it would extend the utility of glycosyltransferases as reagents for glycosylation of complex molecules. Altered glycosyltransferases could also be useful for remodeling cell surfaces and for probing the biological roles of particular carbohydrate structures.